

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
18 January 2001 (18.01.2001)

PCT

(10) International Publication Number  
**WO 01/04790 A1**

(51) International Patent Classification<sup>7</sup>: **G06F 17/20,**  
G10L 13/08, G06K 9/46

(21) International Application Number: PCT/IL00/00399

(22) International Filing Date: 6 July 2000 (06.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
130847 8 July 1999 (08.07.1999) IL

(71) Applicants and

(72) Inventors: **URBACH, Shlomo** [IL/IL]; 43C Gordon Street, 76287 Rehovot (IL). **GIL, Jacob** [IL/IL]; 63/9 Mendes St., 52653 Ramat Gan (IL).

(74) Agent: **FRIEDMAN, Mark, M.**; Beit Samueloff, 7 Hanimanim St., 67897 Tel Aviv (IL).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,

DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

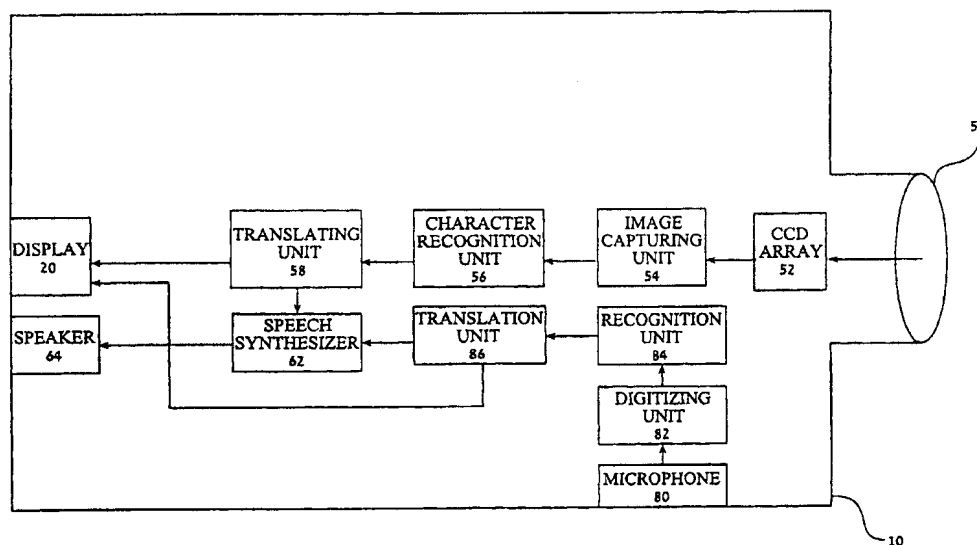
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SIGN TRANSLATOR



(57) Abstract: An automatic device, of particular utility for tourists, for capturing an image at a distance and translating text within the image, comprising: (a) photographing means (54) for capturing an image containing a textual image, (b) character recognition means (56) for converting the textual image into corresponding text data written in a source language, and (c) translation means (58) for translating the text data from the source language to a target language text.

WO 01/04790 A1

## SIGN TRANSLATOR

FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to a system for automatic translation of  
5 text, and, more particularly, to a system for capturing text from a distance and  
automatically providing a text translation into the desired language.

Known in the art are various types of scanners that are capable of  
identifying text and translating the text into another language. The input data  
scanned are converted into system data and input text data using optical  
10 character recognition (OCR) programs, which are widely available and the  
operation of which is well known in the field. OCR programs are often  
designed towards scanning and recognizing alphanumeric text. Some OCR  
programs are designed towards scanning and recognizing ideographic  
characters, such as Japanese kana.

15 An example of one popular commercial application is Quicktionary™,  
a portable, handheld scanner that scans and automatically displays a  
translation. The built-in OCR program is capable of reading and identifying  
multiple fonts and type sizes. The internal memory contains several hundred  
thousand words and idioms.

20 The scanning translators of the known art operate by pressing the paper  
or surface bearing the text (or characters) against the scanning surface of the

scanner. In portable scanning translators, the entire scanner is brought near to the text-bearing surface at which point the scanning is initiated.

Known translation systems are incapable, however, of reading and translating text or characters from afar. This is a serious disadvantage, as often the text-bearing surface is not under thumb and may be substantially inaccessible. For example, millions of tourists each year travel in countries in which for them, the local alphabet, including traffic signs, street signs, advertisements, is incomprehensible. Although in theory it is possible to carry hand-held portable translators and type in the words or sentence that needs translation, in practice such use is clumsy and time-consuming. Moreover, hand-held portable translators often do not have the capability of entering ideographic characters and forms belonging to foreign alphabets such as Cyrillic or Hebrew. Finally, typographical error or improper identification of a foreign form or character by the tourist/user are highly probable, and may often lead to mistranslation or to an inability to provide a translation.

There is therefore a recognized need for, and it would be highly advantageous to have, a system that would be capable of reading a text-bearing and/or character-bearing object from afar and translating the text and/or characters into the requisite language. It would be of particular advantage if such an invention would be economical, reliable, and convenient to use.

SUMMARY OF THE INVENTION

The present invention is an automatic device, of particular utility for tourists, for capturing an image at a distance and translating text within the image, comprising: (a) photographing means for capturing an image  
5 containing a textual image, (b) character recognition means for converting the textual image into corresponding text data written in a source language, and (c) translation means for translating the text data from the source language to a target language text.

According to further features in the described preferred embodiments,  
10 the target language text is provided as visual text.

According to still further features in the described preferred embodiments, the visual text is provided on a visual display.

According to still further features in the described preferred embodiments, the target language text is provided audially.

15 According to still further features in the described preferred embodiments, the photographing device is a digital video camera.

According to still further features in the described preferred embodiments, the photographing device is a digital snapshot (still-photo) camera.

20 According to still further features in the described preferred embodiments, the translation means for translating the text data from the source language to a target language text is provided by accessing the Internet.

According to still further features in the described preferred

embodiments, the accessing of the Internet is performed using a portable computer.

According to still further features in the described preferred embodiments, the accessing of the Internet is performed using a Personal  
5 Digital Assistant (PDA).

According to still further features in the described preferred embodiments, the accessing of the Internet is performed using a mobile phone.

According to still further features in the described preferred embodiments, the visual display is a visual display of a mobile phone.

10 According to still further features in the described preferred embodiments, the automatic device of the present invention further comprises (d) at least one feature selected from the group consisting of: tour books, business information, language-learning means, maps, and travel directions.

According to still further features in the described preferred  
15 embodiments, the automatic device of the present invention further comprises (d) user-identification means.

According to still further features in the described preferred embodiments, the automatic device of the present invention further comprises (d) at least one feature selected from the group consisting of: credit cards,  
20 travel tickets, membership cards, car-rental contracts, hotel reservations, and entertainment tickets.

According to still further features in the described preferred

embodiments, the automatic device of the present invention further comprises (d) an audio device for recording voice data in the source language, and (e) translation means for translating the voice data from the source language to the target language text.

5        According to still further features in the described preferred embodiments, the target language text is provided on a visual display or audially.

         The present invention successfully addresses the shortcomings of the existing technologies by providing a device capable of instantly capturing an  
10    image viewed from a distance, identifying the text therein, and translating the text to the target language.

#### BRIEF DESCRIPTION OF THE DRAWINGS

         The invention is herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to the  
15    drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this  
20    regard, no attempt is made to show structural details of the invention in more detail than is necessary for a fundamental understanding of the invention, the

description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

In the drawings:

FIG. 1a is a schematic illustration of the exterior of a device according to the present invention, integrated into a digital video camera.

FIG. 1b is a schematic illustration of the inner workings of the device of FIG. 1a.

FIG. 2 illustrates the camera display of the system of FIG. 1.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is an automatic system that captures text, usually at a distance, identifies the text and translates the text into the requisite language.

The principles and operation of the device according to the present invention may be better understood with reference to the drawings and the accompanying description.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following description or illustrated in the drawing. The invention is capable of other embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

Referring now to the drawings, Figure 1a is a schematic illustration of the exterior of a device according to the present invention, integrated into a conventional digital video camera 10. The digital video camera 10 has two main operating buttons: a first operating button 12 for filming, and a second operating button 14 for the purpose of the present invention. When the second operating button 14 is pressed, the text is captured, identified, and translated into the requisite (target) language.

Figure 1b is a schematic illustration of the inner workings of the device of Figure 1a, integrated into a conventional digital video camera 10. The lens 50, which represents the optics of the device, focusses a text-containing image (not shown) on to a CCD (Charge Coupled Device) array 52. The voltage of each element of the CCD array corresponds to the light intensity focussed on the element. The image capturing unit 54 reads the voltage of each element in the array, yielding a field of pixels.

The text within the captured image from the image capturing unit 54 is passed on to the character recognition unit 56, which identifies the text. Subsequently, the identified text is passed on to the translating unit 58, which translates the source text into the target language. The translated (target language) text is then provided visually on a display 20. Alternatively, the translated text is passed on to a speech synthesizer 62 which converts the text into audible speech, which is then presented to the user by means of a speaker 64.



It should be emphasized that the system of the present invention can utilize and integrate existing technologies such as OCR, digital video equipment, and computerized translators such as Quicktionary™. While some OCR technologies require that the source language be identified by the user, others are capable of identifying the source language. The present invention can utilize both of these OCR technologies, as well as other kinds of character recognition technologies. In any event, the target language can be dedicated to a particular target language, loaded as a sole target language by the user, or loaded as one of several target languages by the user (in which case, the user must specify which target language is desired). All of these translating options are possible using translating components of the prior art.

Figure 2 illustrates the camera display 20 of the system of Figure 1. Within the camera display 20, a sign 22 having the text 24 "MARKET" is shown. The sign 22 containing text 24 lies within the field of translation 26. Upon pressing the "translate" button (the second operating button 14 of Figure 1a), the text is translated into the desired target language, in this case, Hebrew ( "שווק" ). The translated text 28 is then provided at the bottom 30 of the camera display 20.

In a preferred embodiment, the field of translation 26 is a sub-frame of the camera display 20, preferably located near the center of the camera display 20. In another preferred embodiment, the field of translation 26 can be adjusted according to the size of the text 24.

In the figures and accompanying description above, the device according to the present invention is integrated in a digital camera, such as a digital video camera or a digital snapshot (still-photo) camera. Thus, the inventive features are built into the digital camera. In principle, such a system can include any equipment used to read or obtain visual information at a distance, e.g., digital binoculars or surveying instruments of various kinds.

The present invention may also be an add-on unit to an existing digital camera, i.e., a camera that was manufactured without having the present invention in mind. In connecting the device to an existing camera unit, a short cable having an appropriate fitting is used, the fitting being connected to an IEEE 1394 outlet or similar outlet. Such outlets are typically used today to connect a digital video camera to a computer. Connection to a digital camera is typically effected via existing-type communication ports such as USB.

In add-on systems of this type, wherein the device of the present invention is hooked up to existing equipment (compatible, but not embedded or built-in), the inventive device typically includes a display screen, keyboard (having one or more keys or buttons), and computer communication cable.

In another kind of add-on unit, the device according to the present invention, having a dedicated element for capturing images containing text, is

added-on to a mobile phone or to a Personal Digital Assistant (PDA) such as Palm Pilot™.

Alternatively, the present invention may be a stand-alone unit, having a dedicated element for capturing images containing text instead of being built  
5 into a digital camera.

It must be emphasized that tourists are accustomed to traveling with cameras of various kinds. Because the device according to the present invention can be affixed to cameras and the like, and can even be built as an integral part of the camera, the portable text-capturing and text-translating  
10 device is particularly easy and convenient to use.

The dictionary or dictionaries providing the translation can be stored physically within the device. In a preferred embodiment, the dictionary or dictionaries providing the translation can be accessed via the Internet. This is particularly relevant for applications in which the equipment does not  
15 necessarily have considerable computing power, as with many cellular phones.

As used herein in the specification and in the claims section that follows, the term "text" includes sentences, phrases, words, letters and ideographic signs. The letters include, by way of example, Braille letters,  
20 and the ideographic signs include, by way of example, traffic signs and the like.

As used herein in the specification and in the claims section that follows, the term "translate" refers to any of the following definitions taken

from "The American Heritage College Dictionary" (Houghton Mifflin Co., U.S.A., Third Edition, 1997):

1. To render in another language;
- 2a. To put into simpler terms; explain;
- 5 2b. To express in different words; paraphrase.

The above-mentioned definitions of the term "translate" refer also to ideographic signs, such as Japanese kana, and to traffic signs and the like. The term "translate" also refers to phonetic translation, i.e., a phonetic representation in the source language or in the target language.

10 As used herein in the specification and in the claims section that follows, the term "photographing means" refers to means that are capable of capturing an image at a distance, and that are characteristically capable of capturing the entire image in an instantaneous or all-at-once fashion. Photographing means of the present invention include all kinds of digital  
15 means of capturing an image, including digital cameras (such as digital video cameras and digital snapshot cameras), and digital binoculars. Specifically excluded from photographing means are scanners, in which the image-recording is performed in a gradual fashion, and which require the object bearing the image to be substantially in contact with the scanning  
20 mechanism.

According to still further features in the described preferred embodiments, the automatic device of the present invention further comprises

an audio device for capturing voice data in the source language, and translation means for translating the voice data from the source language to the target language text. Referring to Figure 1b, the audio input is conveyed via microphone 80 to an audio signal digitizing unit 82. The signals are  
5 recognized as words in the recognition unit 84. The words identified in the recognition unit 84 are subsequently translated into the target language by translation unit 86. The translated (target language) text is then provided visually on a display 20. Alternatively, the translated text is passed on to a speech synthesizer 62 which converts the text into audible speech, which is  
10 then presented to the user by means of a speaker 64.

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that  
15 fall within the spirit and broad scope of the appended claims.

## WHAT IS CLAIMED IS:

1. An automatic device for capturing an image at a distance and translating text within the image comprising:

- (a) photographing means for capturing an image containing a textual image;
- (b) character recognition means for converting said textual image into corresponding text data written in a source language;
- (c) translation means for translating said text data from said source language to a target language text.

2. The device of claim 1, wherein said target language text is provided as visual text.

3. The device of claim 2, wherein said visual text is provided on a visual display.

4. The device of claim 1, wherein said target language text is provided audially.

5. The device of claim 1, wherein said photographing device is a digital video camera.

6. The device of claim 1, wherein said photographing device is a

digital snapshot camera.

7. The device of claim 1, wherein said translation means for translating said text data from said source language to a target language text is provided by accessing the Internet.

8. The device of claim 7, wherein said accessing the Internet is performed using a Personal Digital Assistant (PDA).

9. The device of claim 7, wherein said accessing the Internet is performed using a portable computer.

10. The device of claim 7, wherein said accessing the Internet is performed using a mobile phone.

11. The device of claim 3, wherein said visual display is a visual display of a mobile phone.

12. The device of claim 1, further comprising:

(d) at least one feature selected from the group consisting of tour book, business information, language-learning means, map, and travel directions.

13. The device of claim 1, further comprising:

(d) user-identification means.

14. The device of claim 1, further comprising:

(d) at least one feature selected from the group consisting of credit card, travel ticket, membership card, car-rental contract, hotel reservation, and entertainment ticket.

15. The device of claim 1, further comprising:

(d) an audio device for capturing voice data in said source language, and

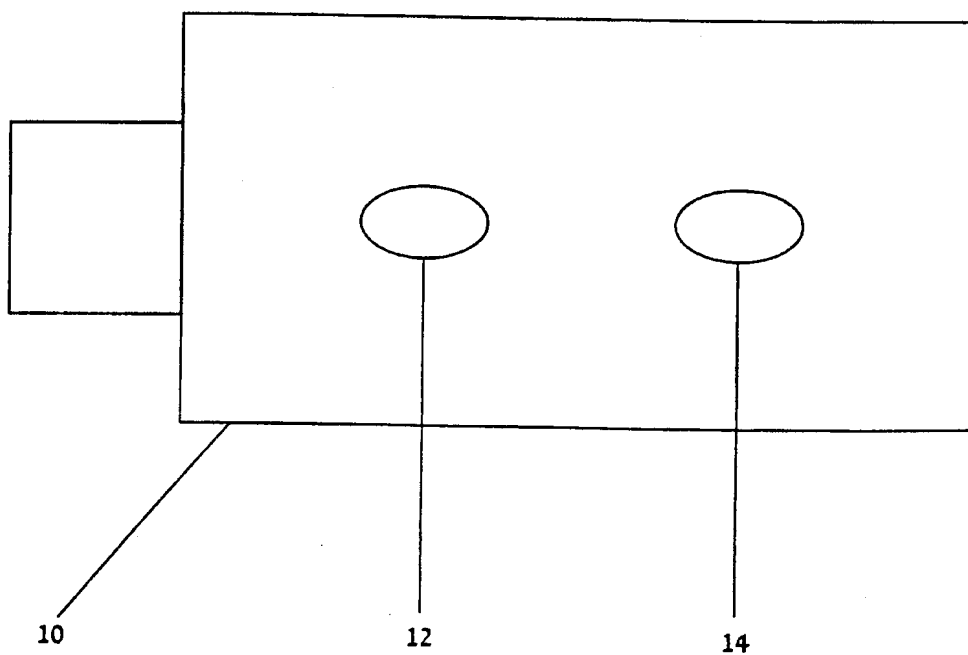
(e) translation means for translating said voice data from said source language to said target language text.

16. The device of claim 15, wherein said target language text is provided on a visual display.

17. The device of claim 15, wherein said target language text is provided audially.



1/3



**FIG.1a**

**SUBSTITUTE SHEET (RULE 26)**

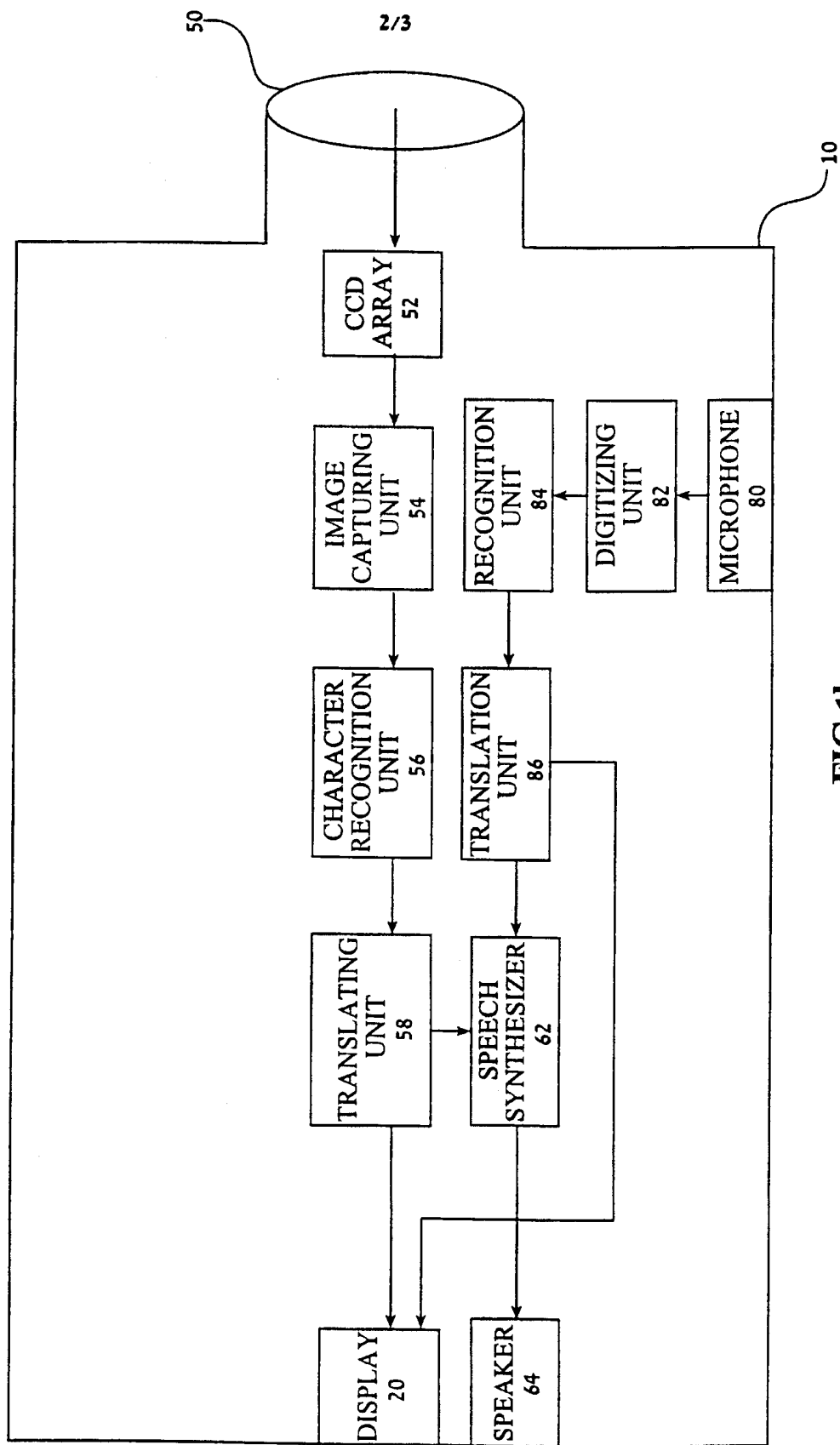
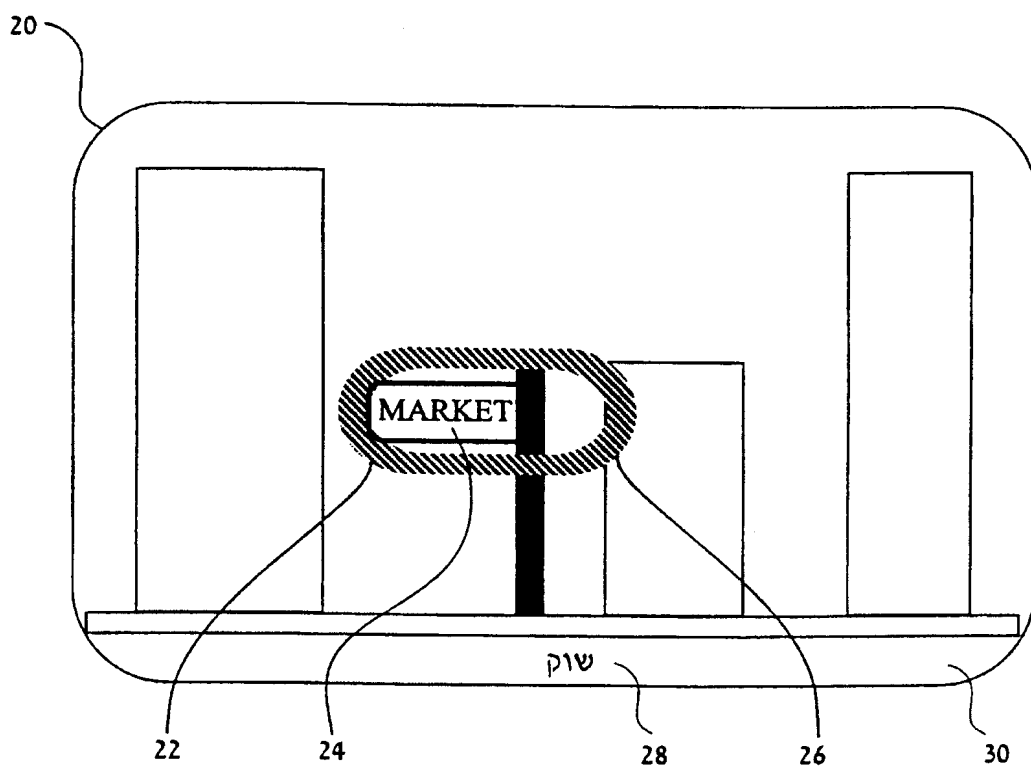


FIG. 1b



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/IL00/00399

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/20; G10L 13/08; G06K 9/46

US CL : 704/1, 277; 382/190

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 704/1, 9, 10, 270, 271, 272, 277; 382/190, 229

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y ----A	US 5,884,249 A (NAMBA et al.) 16 March 1999, abstract; figs. 1-2, 8, & 12-18; col. 2, line 1 to col. 3, line 23; col. 5, line 35 to col. 6, line 36; col. 13, line 1 to col. 15, line 38; col. 18, line 46 to col. 24, line 47; and col. 26, line 50 to col. 28, line 27	7-9 -----1-6 & 10-17
A	US 5,901,246 A (HOFFBERG et al) 04 May 1999, abstract; col. 73, lines 36-62; col. 75, line 38 to col. 86, line 34; col. 88, line 33 to col. 89, line 25; and col. 95, lines 1-34	1-17

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

03 NOVEMBER 2000

Date of mailing of the international search report

06 DEC 2000

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JOSEPH THOMAS

Telephone No. (703) 308-3900

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/IL00/00399

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	US 5,917,944 A (WAKISAKA, et al) 29 June 1999, abstract, figs. 1, 6A-6B, 8, 12A-12B, & 13; col. 2, line 42 to col. 6, line 8; col. 6, line 54 to col. 9, line 23; col. 10, line 52 to col. 15, line 53; and col. 17, line 12 to col. 18, line 58	1-6 & 12-17 ----- 7-11
Y, P ----- A, P	US 5,956,681 A (YAMAKITA) 21 September 1999, abstract, figs. 1-3; col. 1, line 7 to col. 2, line 50; col. 3, line 29 to col. 9, line 45; and col. 35, line 32 to col. 36, line 36	10-11 ----- 1-9 & 12-17

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/IL00/00399

### B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

#### EAST Search

search terms: translation/interpretation, OCR or optical character, photograph or camera, PDA or personal/digital assistants